AD-A092 636   STANFORD UNIV CA DEPT OF STATISTICS   F/G 12/1

COCHRAN'S THEOREM, RANK ADDITIVITY, AND TRIPOTENT MATRICES.(U)

AUG 80 T W ANDERSON, G P STYAN   N00014-75-C-0442

UNCLASSIFIED TR-43   NL

1 OF 1

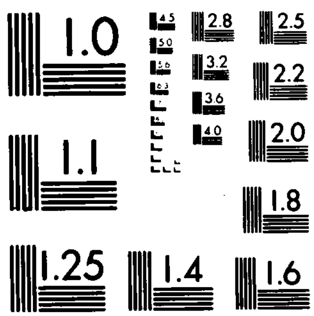
AD A
092636

END

DATE FILMED

1-81

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

COCHRAN'S THEOREM, RANK ADDITIVITY, AND TRIPOTENT MATRICES

BY

T. W. ANDERSON and GEORGE P. H. STYAN

TECHNICAL REPORT NO. 43

AUGUST 1980

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

COCHRAN'S THEOREM, RANK ADDITIVITY,
AND TRIPOTENT MATRICES

by

T. W. Anderson
Stanford University

and

George P. H. Styan
McGill University

TECHNICAL REPORT NO. 43
AUGUST 1980

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)

OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Cochran's Theorem, Rank Additivity,

and Tripotent Matrices

T. W. Anderson and George P. H. Styan

Stanford University and McGill University

1. Introduction

Let $\underset{\sim}{x}$ be a $p \times 1$ random vector distributed according to a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{I}_p$. We will denote this by $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{0}, \underset{\sim}{I}_p)$. Let $q_1, \ldots, q_k$ be quadratic forms in $\underset{\sim}{x}$ with ranks $r_1, \ldots, r_k$, respectively, and suppose that $\sum q_i = \underset{\sim}{x}'\underset{\sim}{x}$. Then what has become well known as Cochran's Theorem is Theorem II of Cochran (1934, p. 179): <u>A necessary and sufficient condition that</u> $q_1, \ldots, q_k$ <u>be independently distributed as</u> χ^2 <u>is that</u> $\sum r_i = p$.

Rao (1973, §3b.4) gives this result with $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{\mu}, \underset{\sim}{I})$ as the Fisher-Cochran Theorem. Fisher (1925) showed that if the quadratic form q in $\underset{\sim}{x}$ is distributed as χ^2_h then $\underset{\sim}{x}'\underset{\sim}{x} - q$ is distributed as χ^2_{p-h} independently of q, cf. James (1952).

Our purpose in this paper is to review various extensions of Cochran's Theorem in a bibliographic and historical perspective, with special emphasis on matrix-theoretic analogues. While we present over 30 references, we note that Scarowsky (1973) has a rather complete discussion and bibliography on the distribution of quadratic forms in

normal random variables. See also the bibliography by Anderson, Das Gupta, and Styan (1972), where 90 research papers published through 1966 are listed under subject-matter code 2.5 (distribution of quadratic and bilinear forms in normal variables).

The first section is devoted to a survey of results summarized in Theorems 1.1 and 1.2. The proofs are given in Section 2. In the following section the extensions from idempotent to tripotent matrices are given and proved.

To formulate our first matrix-theoretic extension of Cochran's Theorem we let $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_k$ be p×p symmetric matrices and write $\underset{\sim}{A} = \sum \underset{\sim}{A}_i$. Consider the following statements:

(a) $\underset{\sim}{A}_i^2 = \underset{\sim}{A}_i$, $\qquad i=1,\ldots,k$,

(b) $\underset{\sim}{A}_i \underset{\sim}{A}_j = \underset{\sim}{0}$ $\qquad$ for all $i \neq j$,

(c) $\underset{\sim}{A} = \underset{\sim}{I}$,

(d) $\sum \text{rank}(\underset{\sim}{A}_i) = \text{rank}(\underset{\sim}{A})$.

Then the matrix-theoretic analogue of Cochran's Theorem is:

$$(a), (b), (c) \rightarrow (d), \tag{1.1}$$

$$(c), (d) \rightarrow (a), (b). \tag{1.2}$$

The reason that these two versions of Cochran's Theorem are equivalent follows from the following two well-known results:

LEMMA 1.1. *Let* $x \sim N(\mu, \Sigma)$, *with* Σ *positive definite, and let* A *be nonrandom and symmetric. Then* $x'Ax \sim \chi_f^2(\delta^2)$, *a noncentral* χ^2 *distribution with* f *degrees of freedom and noncentrality parameter* δ^2, *if and only if* $A\Sigma A = A$, *and then* $f = \mathrm{tr}A = \mathrm{rank}(A)$ *and* $\delta^2 = \mu' A \mu$.

We write $\mathrm{tr}A$ for the trace of A and note that when $A = A^2$ then $\mathrm{tr}A = \mathrm{rank}(A)$; this result holds even when A is not symmetric (cf., e.g., Rao (1973), p. 28).

When $\Sigma = I$ the condition in Lemma 1.1 reduces to $A^2 = A$, and this was first given by Craig (1943) with $\mu = 0$ and then by Carpenter (1950) with μ possibly nonzero. (Thus (a) is equivalent to $q_i = x'A_i x$ having a χ^2 distribution with number of degrees of freedom equal to $\mathrm{rank}(A_i)$.) Sakamoto (1944, Th. II, p. 5) gave the more general version, with Σ positive definite and $\mu = 0$. Cochran (1934, Corollary 1, p. 179) took $x \sim N(0, I)$ and gave Lemma 1.1 with the condition that all the nonzero eigenvalues of A be equal to 1 instead of the condition $A^2 = A$.

LEMMA 1.2. *Let* x *and* A *be defined as in Lemma 1.1 and let* B *be nonrandom and symmetric. Then* $x'Ax$ *and* $x'Bx$ *are independently distributed if and only if* $A\Sigma B = 0$.

When $\Sigma = I$ the condition in Lemma 1.2 reduces to $AB = 0$, and this was first given by Craig (1943) with $\mu = 0$ and then by Carpenter (1950) with μ possibly nonzero. Again Sakamoto (1944, Th. I, p. 5) gave the more general version with Σ positive definite and $\mu = 0$. Their proofs, however, turned out to be incorrect and the first correct proof of Lemma 1.2 (with $\mu = 0$) seems to be by Ogawa (1948; 1949, cf. p. 85). Cochran (1934, Theorem III, p. 181) let $x \sim N(0, I)$ and gave the condition in Lemma 1.2 as

$$|\underset{\sim}{I} - is\underset{\sim}{A}| \cdot |\underset{\sim}{I} - it\underset{\sim}{B}| = |\underset{\sim}{I} - is\underset{\sim}{A} - it\underset{\sim}{B}| \qquad (1.3)$$

for all real s and t, where $i = \sqrt{-1}$ and $|\cdot|$ denotes determinant. Ogasawara and Takahashi (1951, Lemma 1) gave a short proof that (1.3) implies $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{0}$ when the symmetric matrice $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are not necessarily positive semi-definite.

Cochran's Theorem was first extended to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{\mu}, \underset{\sim}{I}_p)$ by Madow (1940) and then to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, $\underset{\sim}{\Sigma}$ positive definite, by Ogawa (1946, 1947), who also relaxed the condition (c) to $\underset{\sim}{A}^2 = \underset{\sim}{A}$. Ogasawara and Takahashi (1951) extended Cochran's Theorem to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$, $\underset{\sim}{\Sigma}$ positive definite, and to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, with $\underset{\sim}{\Sigma}$ possibly singular. Extensions to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$, with $\underset{\sim}{\Sigma}$ possibly singular, have been given by Styan (1970, Theorem 6) and Tan (1977, Theorem 4.2); Ogasawara and Takahashi (1951) extended Lemmas 1.1 and 1.2 to $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$, with $\underset{\sim}{\Sigma}$ possibly singular.

James (1952) appears to be the first to notice that (1.1) could be extended to

$$\text{(a), (c)} \rightarrow \text{(b), (d),}$$

$$\text{(b), (c)} \rightarrow \text{(a), (d),}$$

while

$$\text{(a), (b)} \rightarrow \underset{\sim}{A}^2 = \underset{\sim}{A} \text{ and (d)}$$

follows at once from the definition of the χ^2-distribution.

Chipman and Rao (1964) and Khatri (1968) extended the matrix analogue of Cochran's Theorem to square matrices which are not necessarily symmetric:

THEOREM 1.1. Let $A_1, \ldots, A_k$ be square matrices, not necessarily symmetric, and let $A = \sum A_i$. Consider the following statements:

$$\text{(a)} \qquad A_i^2 = A_i, \qquad\qquad i=1,\ldots,k,$$

$$\text{(b)} \qquad A_i A_j = 0 \qquad\qquad \text{for all } \; i \neq j,$$

$$\text{(c)} \qquad A^2 = A,$$

$$\text{(d)} \qquad \sum \text{rank}(A_i) = \text{rank}(A),$$

$$\text{(e)} \qquad \text{rank}(A_i^2) = \text{rank}(A_i), \qquad\qquad i=1,\ldots,k.$$

Then

$$\text{(a), (b)} \rightarrow \text{(c), (d), (e),} \tag{1.4}$$

$$\text{(a), (c)} \rightarrow \text{(b), (d), (e),} \tag{1.5}$$

$$\text{(b), (c), (e)} \rightarrow \text{(a), (d),} \tag{1.6}$$

$$\text{(c), (d)} \rightarrow \text{(a), (b), (e).} \tag{1.7}$$

As Rao and Mitra (1971, p. 112) point out, the extra condition (e) in (1.6) is required (to cover possible asymmetry); for if k=2 and

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},$$

then (b), (c) hold, but (a) and (d) do not. Banerjee and Nagase (1976) replace the extra condition (e) in (1.6) by

$$\text{(f)} \qquad \text{rank}(A_i) = \text{tr}A_i, \qquad\qquad i=1,\ldots,k,$$

and prove that

$$(b), (c), (f) \rightarrow (a), (d); \tag{1.8}$$

however, the condition (b) is now no longer required on the left of (1.8) since

$$(c), (f) \rightarrow (a), (b), (d)$$

follows from

$$\operatorname{rank}(\underset{\sim}{A}) = \operatorname{tr}\underset{\sim}{A} = \operatorname{tr}\sum \underset{\sim}{A}_i = \sum \operatorname{tr}\underset{\sim}{A}_i = \sum \operatorname{rank}(\underset{\sim}{A}_i)$$

and (1.7).

In Section 2 we present several proofs of Theorem 1.1.

Marsaglia and Styan (1974) extended Theorem 1.1 by considering an arbitrary sum of matrices, which may now be rectangular. The analogue of Theorem 1.1 is

THEOREM 1.2. *Let* $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_k$ *be* $p \times q$ *matrices, and let* $\underset{\sim}{A} = \sum \underset{\sim}{A}_i$. *Consider the following statements*:

$$\text{(a')} \qquad \underset{\sim}{A}_i \underset{\sim}{A}^- \underset{\sim}{A}_i = \underset{\sim}{A}_i, \qquad i=1,\ldots,k,$$

$$\text{(b')} \qquad \underset{\sim}{A}_i \underset{\sim}{A}^- \underset{\sim}{A}_j = \underset{\sim}{0} \qquad \textit{for all } i \neq j,$$

$$\text{(c')} \qquad \operatorname{rank}(\underset{\sim}{A}_i \underset{\sim}{A}^- \underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}_i), \qquad i=1,\ldots,k,$$

$$\text{(d')} \qquad \sum \operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}),$$

where $\underset{\sim}{A}^-$ *is some g-inverse of* $\underset{\sim}{A}$. *Then*

$$(a') \rightarrow (b'), (c'), (d'), \tag{1.9}$$

$$(b'), (c') \rightarrow (a'), (d'), \tag{1.10}$$

$$(d') \rightarrow (a'), (b'), (c'). \tag{1.11}$$

If (a') or if (b') and (c') hold for some g-inverse $\underset{\sim}{A}^-$ then (a'), (b') and (c') hold for every g-inverse $\underset{\sim}{A}^-$.

In Theorem 1.2 we define a g-inverse of $\underset{\sim}{A}$ as any solution $\underset{\sim}{A}^-$ to $\underset{\sim}{A}\underset{\sim}{A}^-\underset{\sim}{A} = \underset{\sim}{A}$, cf. Rao (1962), Rao and Mitra (1971).

The condition (c') in Theorem 1.2 plays the role of condition (e) in Theorem 1.1.

Marsaglia and Styan (1974, Th. 13) proved (1.11), while Hartwig (1980) has established (1.9). The proposition (1.10), however, appears to be new and is proved in Section 2, where we also present several different proofs of (1.7). In Section 3 we extend Theorem 1.1 to tripotent matrices, following the work by Luther (1965), Tan (1975, 1976) and Khatri (1977). In Section 4 we discuss the applications of these algebraic theorems to statistics.

2. Some Proofs

2.1. Proof of Theorem 1.1. To prove (1.7) in Theorem 1.1 we begin by reducing condition (c) to a sum being I as in the earlier version of Cochran's Theorem; then (1.7) reduces to (1.2). We may do this since if $\underset{\sim}{A}$ is p×p, not necessarily symmetric, then, as we shall show,

$$\underset{\sim}{A}^2 = \underset{\sim}{A} \leftrightarrow \operatorname{rank}(\underset{\sim}{I} - \underset{\sim}{A}) = p - \operatorname{rank}(\underset{\sim}{A}). \tag{2.1}$$

(Note Fisher's 1925 result goes both ways, cf. Section 1, paragraph 2.) To prove (2.1) let $\underset{\sim}{A}^2 = \underset{\sim}{A}$; then $(\underset{\sim}{I} - \underset{\sim}{A})^2 = \underset{\sim}{I} - \underset{\sim}{A}$ and so

$$\operatorname{rank}(\underset{\sim}{I} - \underset{\sim}{A}) = \operatorname{tr}(\underset{\sim}{I} - \underset{\sim}{A}) = p - \operatorname{tr}\underset{\sim}{A} = p - \operatorname{rank}(\underset{\sim}{A}).$$

To go the other way we follow Krafft (1978, pp. 407-408) by noting that

$$\mathcal{N}(\mathbf{A}) \subset \mathcal{C}(\mathbf{I} - \mathbf{A}), \tag{2.2}$$

where $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}$ is the null space of $\mathbf{A}$ and $\mathcal{C}(\mathbf{I} - \mathbf{A}) = \{(\mathbf{I} - \mathbf{A})\mathbf{x}\}$ is the column space of $\mathbf{I} - \mathbf{A}$. (If $\mathbf{x} \in \mathcal{N}(\mathbf{A})$, then $\mathbf{A}\mathbf{x} = \mathbf{0}$ and $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{x} \in \mathcal{C}(\mathbf{I} - \mathbf{A})$.) If $\mathrm{rank}(\mathbf{I} - \mathbf{A}) = p - \mathrm{rank}(\mathbf{A})$, then equality must hold in (2.2) and so $\mathbf{A}^2 = \mathbf{A}$.

We now write $\mathbf{A}_0 = \mathbf{I} - \mathbf{A}$, and in view of (2.1) we replace (c) by $\sum_{i=0}^{p} \mathbf{A}_i = \mathbf{I}$, and (d) by $\sum_{i=0}^{k} \mathrm{rank}(\mathbf{A}_i) = p$.

The proof of (1.7) by Cochran (1934, p. 180), cf. also Anderson (1958, p. 164) and Rao (1973, §3b.4), requires that $\mathbf{A}_1, \ldots, \mathbf{A}_k$ be symmetric. In this event we may write

$$\mathbf{A}_i = \mathbf{P}_i\mathbf{P}_i' - \mathbf{Q}_i\mathbf{Q}_i', \qquad i=0,1,\ldots,k, \tag{2.3}$$

where $\mathbf{P}_i$ is $p \times p_i$, $\mathbf{Q}_i$ is $p \times q_i$, and $\mathbf{A}_i$ has p_i positive and q_i negative eigenvalues, cf. e.g., Anderson (1958, p. 346). In (2.3) we assume that $\mathbf{P}_i$ has rank p_i, $\mathbf{Q}_i$ has rank q_i, and $p_i + q_i = r_i$, the rank of $\mathbf{A}_i$. Hence

$$\begin{aligned}
\mathbf{I}_p &= \sum_{i=0}^{k} \mathbf{A}_i = \sum_{i=0}^{k} \mathbf{P}_i\mathbf{P}_i' - \sum_{i=0}^{k} \mathbf{Q}_i\mathbf{Q}_i' \\
&= (\mathbf{P}_0, \ldots, \mathbf{P}_k, \mathbf{Q}_0, \ldots, \mathbf{Q}_k) \begin{pmatrix} \mathbf{I}_{p-q} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_q \end{pmatrix} \begin{pmatrix} \mathbf{P}_0' \\ \vdots \\ \mathbf{P}_k' \\ \mathbf{Q}_0' \\ \vdots \\ \mathbf{Q}_k' \end{pmatrix} \\
&= \mathbf{P}\mathbf{J}\mathbf{P}',
\end{aligned} \tag{2.4}$$

say, where $q = \sum_{i=0}^{k} q_i$, since from (d) now $p = \sum_{i=0}^{k} r_i = \sum_{i=0}^{k}(p_i + q_i) = (\sum_{i=0}^{k} p_i) + q$. But (2.4) is positive definite and $\mathbf{P}$ is nonsingular;

hence $q=0$ and $\mathbf{J}=\mathbf{I}_p$. Thus $q_0 = \ldots = q_k = 0$ and (2.4) reduces to

$$\mathbf{I}_p = (\mathbf{P}_0,\ldots,\mathbf{P}_k)\begin{pmatrix} \mathbf{P}_0' \\ \vdots \\ \mathbf{P}_k' \end{pmatrix} = \mathbf{P}\mathbf{P}',$$

and so $\mathbf{P} = (\mathbf{P}_0,\ldots,\mathbf{P}_k)$ is an orthogonal matrix. Hence $\mathbf{A}_i^2 = \mathbf{P}_i\mathbf{P}_i'\mathbf{P}_i\mathbf{P}_i' = \mathbf{P}_i\mathbf{P}_i' = \mathbf{A}_i$ since $\mathbf{P}_i'\mathbf{P}_i = \mathbf{I}_{r_i}$, and $\mathbf{A}_i\mathbf{A}_j = \mathbf{P}_i\mathbf{P}_i'\mathbf{P}_j\mathbf{P}_j' = \mathbf{0}$ for all $i\neq j$ since then $\mathbf{P}_i'\mathbf{P}_j = \mathbf{0}$.

We now present three other proofs of (1.7); these three proofs do not require that $\mathbf{A}_0,\ldots,\mathbf{A}_k$ be symmetric.

Following Craig (1938, p. 49), cf. also Aitken (1950, §6) and Rao and Mitra (1971, pp. 111-112), we may prove (1.7) using a rank-subadditivity argument. From (2.1) with $\mathbf{A}_k$ replacing $\mathbf{A}$ we have

$$\begin{aligned} p - \operatorname{rank}(\mathbf{A}_k) &\le \operatorname{rank}(\mathbf{I}_p - \mathbf{A}_k) \\ &= \operatorname{rank}(\mathbf{A}_0 + \ldots + \mathbf{A}_{k-1}) \\ &\le \operatorname{rank}(\mathbf{A}_0) + \ldots + \operatorname{rank}(\mathbf{A}_{k-1}) \\ &= p - \operatorname{rank}(\mathbf{A}_k) \end{aligned} \tag{2.5}$$

when (d) holds. This inequality string, therefore, collapses, and $\operatorname{rank}(\mathbf{I}_p - \mathbf{A}_k) = p - \operatorname{rank}(\mathbf{A}_k)$, which implies $\mathbf{A}_k^2 = \mathbf{A}_k$ by (2.1); repeating the argument with $\mathbf{A}_{k-1}, \mathbf{A}_{k-2},\ldots$ yields (a). To see that this implies (b) we follow Rao and Mitra (1971, p. 112) by noting that the argument used in (2.5) implies that

$$(\mathbf{A}_i + \mathbf{A}_j)^2 = \mathbf{A}_i + \mathbf{A}_j$$

and so

$$\mathbf{A}_i\mathbf{A}_j + \mathbf{A}_j\mathbf{A}_i = \mathbf{0}.$$

Premultiplying by $\underset{\sim}{A}_i$ yields

$$\underset{\sim}{A}_i\underset{\sim}{A}_j + \underset{\sim}{A}_i\underset{\sim}{A}_j\underset{\sim}{A}_i = \underset{\sim}{0}, \tag{2.6}$$

while postmultiplying (2.6) by $\underset{\sim}{A}_i$ yields

$$2\underset{\sim}{A}_i\underset{\sim}{A}_j\underset{\sim}{A}_i = \underset{\sim}{0} = \underset{\sim}{A}_i\underset{\sim}{A}_j\underset{\sim}{A}_i.$$

Substituting into (2.6) yields (b).

Our next proof of (1.7) follows Chipman and Rao (1964, p. 4), cf. also Styan (1970, p. 571). We write

$$\underset{\sim}{A}_i = \underset{\sim}{B}_i\underset{\sim}{C}_i',$$

where $\underset{\sim}{B}_i$ and $\underset{\sim}{C}_i$ are $p \times r_i$ of rank r_i. Then

$$\begin{aligned} \underset{\sim}{I}_p = \sum \underset{\sim}{A}_i &= \sum \underset{\sim}{B}_i\underset{\sim}{C}_i' \\ &= (\underset{\sim}{B}_0, \ldots, \underset{\sim}{B}_k) \begin{pmatrix} \underset{\sim}{C}_0' \\ \vdots \\ \underset{\sim}{C}_k' \end{pmatrix}, \\ &= \underset{\sim}{B}\underset{\sim}{C}', \end{aligned}$$

say. By (d) $\underset{\sim}{B}$ and $\underset{\sim}{C}$ are both nonsingular and so $\underset{\sim}{C}' = \underset{\sim}{B}^{-1}$ and

$$\underset{\sim}{C}'\underset{\sim}{B} = \underset{\sim}{I}_p = \begin{pmatrix} \underset{\sim}{C}_0'\underset{\sim}{B}_0 & , \ldots , & \underset{\sim}{C}_0'\underset{\sim}{B}_k \\ \vdots & & \vdots \\ \underset{\sim}{C}_k'\underset{\sim}{B}_0 & , \ldots , & \underset{\sim}{C}_k'\underset{\sim}{B}_k \end{pmatrix},$$

which implies that

$$\underset{\sim}{A}_i^2 = \underset{\sim}{B}_i\underset{\sim}{C}_i'\underset{\sim}{B}_i\underset{\sim}{C}_i' = \underset{\sim}{B}_i\underset{\sim}{C}_i' = \underset{\sim}{A}_i,$$

$$\underset{\sim}{A}_i\underset{\sim}{A}_j = \underset{\sim}{B}_i\underset{\sim}{C}_i'\underset{\sim}{B}_j\underset{\sim}{C}_j' = \underset{\sim}{0} \qquad \text{for all } i \neq j.$$

Hence (1.7) is established.

Our last proof of (1.7) follows Loynes (1966), cf. also Searle (1971, p. 63). A rank-subadditivity argument is used similar to that used in (2.5):

$$\begin{aligned} p - \operatorname{rank}(\underset{\sim}{A}_k) &\le \operatorname{rank}(\underset{\sim}{I}_p - \underset{\sim}{A}_k) \\ &\le \operatorname{rank}(\underset{\sim}{A}_0, \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_{k-1}, \underset{\sim}{I}_p - \underset{\sim}{A}_k) \\ &= \operatorname{rank}(\underset{\sim}{A}_0, \ldots, \underset{\sim}{A}_{k-1}, \underset{\sim}{I} - \underset{\sim}{A}_0 - \ldots - \underset{\sim}{A}_{k-1} - \underset{\sim}{A}_k) \\ &= \operatorname{rank}(\underset{\sim}{A}_0, \ldots, \underset{\sim}{A}_{k-1}) \\ &\le \operatorname{rank}(\underset{\sim}{A}_0) + \ldots + \operatorname{rank}(\underset{\sim}{A}_{k-1}) \\ &= p - \operatorname{rank}(\underset{\sim}{A}_k). \end{aligned}$$

The rest of Theorem 1.1 is easily proved. To prove (1.4) we see that (a), (b) $\rightarrow$

$$\underset{\sim}{A}^2 = (\sum \underset{\sim}{A}_i)^2 = \sum \underset{\sim}{A}_i^2 + \sum_{i \neq j} \underset{\sim}{A}_i \underset{\sim}{A}_j = \sum \underset{\sim}{A}_i = \underset{\sim}{A},$$

while

$$\sum \operatorname{rank}(\underset{\sim}{A}_i) = \sum \operatorname{tr} \underset{\sim}{A}_i = \operatorname{tr} \sum \underset{\sim}{A}_i = \operatorname{tr} \underset{\sim}{A} = \operatorname{rank}(\underset{\sim}{A}), \tag{2.7}$$

and so (1.4) is established.

To prove (1.5) we see that (a), (c) $\rightarrow$ (d) from (2.7) and so (1.5) follows from (1.7).

To prove (1.6) we see that (b), (c) $\rightarrow$

$$\underset{\sim}{A}^2 = (\sum \underset{\sim}{A}_i)^2 = \sum \underset{\sim}{A}_i^2 + \sum_{i \neq j} \underset{\sim}{A}_i \underset{\sim}{A}_j = \sum \underset{\sim}{A}_i^2 = \sum \underset{\sim}{A}_i = \underset{\sim}{A};$$

multiplying through by $\underset{\sim}{A}_i$ yields

$$\mathbf{A}_i^3 = \mathbf{A}_i^2 \tag{2.8}$$

using (b). To see that (2.8) $\rightarrow$ (a) we use the rank cancellation rule (2.13) in Marsaglia and Styan (1974, p. 271); this rule will also be useful later on.

LEMMA 2.1. Right-hand Rank Cancellation Rule. *If*

$$\mathbf{LAX} = \mathbf{MAX} \text{ and } \operatorname{rank}(\mathbf{AX}) = \operatorname{rank}(\mathbf{A}) \tag{2.9}$$

for some conformable matrices $\mathbf{A}$, $\mathbf{L}$, $\mathbf{M}$ *and* $\mathbf{X}$, *then*

$$\mathbf{LA} = \mathbf{MA}. \tag{2.10}$$

Thus (2.8) $\rightarrow$ (a) by replacing $\mathbf{L}$, $\mathbf{A}$ and $\mathbf{X}$ in (2.9) by $\mathbf{A}_i$ and $\mathbf{M}$ by $\mathbf{I}$. Then (2.9) becomes (2.8) and (e), while (2.10) becomes (a). (We note that the two matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ displayed right after Theorem 1.1 satisfy (2.8) but not (e).) Then (d) follows from (1.4) or (1.5).

Proof of Lemma 2.1. Let $\mathbf{A} = \mathbf{BC}'$, where $\mathbf{B}$ and $\mathbf{C}$ have r columns and $r = \operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{B}) = \operatorname{rank}(\mathbf{C})$. Then $\operatorname{rank}(\mathbf{AX}) = \operatorname{rank}(\mathbf{BC}'\mathbf{X}) = \operatorname{rank}(\mathbf{C}'\mathbf{X}) = \operatorname{rank}(\mathbf{A})$, and so $\mathbf{C}'\mathbf{X}$ has full row rank. Thus $\mathbf{LAX} = \mathbf{MAX}$ equals $\mathbf{LBC}'\mathbf{X} = \mathbf{MBC}'\mathbf{X} \rightarrow \mathbf{LB} = \mathbf{MB} \rightarrow \mathbf{LBC}' = \mathbf{MBC}'$, which is (2.10). Q.E.D.

Transposing the matrices in Lemma 2.1 yields:

LEMMA 2.2. Left-hand Rank Cancellation Rule. *If*

$$\mathbf{LAX} = \mathbf{LAY} \text{ and } \operatorname{rank}(\mathbf{LA}) = \operatorname{rank}(\mathbf{A})$$

for some conformable matrices $\mathbf{A}$, $\mathbf{L}$, $\mathbf{X}$ *and* $\mathbf{Y}$, *then*

$$\mathbf{AX} = \mathbf{AY}.$$

2.2. Proof of Theorem 1.2. Premultiplying (a') by $\underset{\sim}{A}^-$ yields (a) of Theorem 1.1 with $\underset{\sim}{A}_i$ replaced by $\underset{\sim}{A}^-\underset{\sim}{A}_i$. Moreover, condition (c) of Theorem 1.1 now always holds since $\underset{\sim}{A}^-\underset{\sim}{A} = \sum \underset{\sim}{A}^-\underset{\sim}{A}_i$ is always idempotent. Hence (1.5) implies that $\underset{\sim}{A}^-\underset{\sim}{A}_i\underset{\sim}{A}^-\underset{\sim}{A}_j = 0$ for all $i \neq j$. Premultiplying by $\underset{\sim}{A}_i$ and using (a') yields (b'). Furthermore (1.5) implies that

$$\sum \operatorname{rank}(\underset{\sim}{A}^-\underset{\sim}{A}_i) = \operatorname{rank} \sum \underset{\sim}{A}^-\underset{\sim}{A}_i = \operatorname{rank}(\underset{\sim}{A}),$$

which reduces to (d') since (a')

$$\operatorname{rank}(\underset{\sim}{A}^-\underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}_i), \tag{2.11}$$

which follows from

$$\operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}_i\underset{\sim}{A}^-\underset{\sim}{A}_i) \leq \operatorname{rank}(\underset{\sim}{A}^-\underset{\sim}{A}_i) \leq \operatorname{rank}(\underset{\sim}{A}_i).$$

Thus (1.9) is established.

We now prove (1.11). Condition (d') implies

$$\sum \operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}) = \operatorname{rank}(\underset{\sim}{A}^-\underset{\sim}{A}) = \operatorname{rank}(\sum \underset{\sim}{A}^-\underset{\sim}{A}_i) \leq \sum \operatorname{rank}(\underset{\sim}{A}^-\underset{\sim}{A}_i)$$

$$\leq \sum \operatorname{rank}(\underset{\sim}{A}_i) \tag{2.12}$$

and so (d) of Theorem 1.1 holds with $\underset{\sim}{A}_i$ replaced by $\underset{\sim}{A}^-\underset{\sim}{A}_i$. Since (c) now always holds we obtain in lieu of (a)

$$\underset{\sim}{A}^-\underset{\sim}{A}_i\underset{\sim}{A}^-\underset{\sim}{A}_i = \underset{\sim}{A}^-\underset{\sim}{A}_i \tag{2.13}$$

by (1.7). But (2.12) implies (2.11) and so we may cancel the front $\underset{\sim}{A}^-$ on both sides of (2.12) using Lemma 2.2 to yield (a'). The rest of (1.11) follows from (1.9).

To prove (1.10) we use the same technique which we used above to yield

$$A^- A_i A^- A_i A^- A_i = A^- A_i A^- A_i, \tag{2.14}$$

which is (2.8) with A_i replaced by $A^- A_i$. The rank condition (c') and Lemma 2.1 allow us to cancel the $A^- A_i$ on the right of both sides of (2.14) to yield

$$A^- A_i A^- A_i = A^- A_i. \tag{2.15}$$

Using (2.11) and Lemma 2.2 allows us to cancel the leading A^- on both sides of (2.15) and this yields (a'). The rest of (1.10) follows from (1.9) and the proof is complete. Q.E.D.

We may extend Theorem 1.1 to tripotent matrices using Theorem 1.2. We do this in the next section.

3. Tripotent Matrices

A square matrix A is said to be *tripotent* whenever $A^3 = A$. Tripotent matrices have been studied by Luther (1965), Tan (1975, 1976) and Khatri (1977). These authors considered extending Theorem 1.1 to $A_1, \ldots, A_k$ tripotent. This is of interest in statistics since if $x \sim \underline{N}(0, I)$ and A is symmetric nonrandom then $x'Ax$ is distributed as the difference of two independently distributed χ^2-variates if and only if $A^3 = A$, cf. Graybill (1969, p. 352), Tan (1975, Theorem 3.5).

Consider, therefore, the following statements:

(a") $\quad \underset{\sim}{A}_i^3 = \underset{\sim}{A}_i, \qquad i=1,\ldots,k,$

(b") $\quad \underset{\sim}{A}_i\underset{\sim}{A}_j = \underset{\sim}{0} \qquad$ for all $i\neq j$,

(c") $\quad \underset{\sim}{A}^3 = \underset{\sim}{A},$

(d") $\sum\text{rank}(\underset{\sim}{A}_i) = \text{rank}(\underset{\sim}{A}).$

Then it is easy to see that (a"), (b") $\rightarrow$ (c"). To see that (d") is also implied we note that when $\underset{\sim}{A}^3 = \underset{\sim}{A}$ then, cf. Graybill (1969, Theorem 12.4.4), Rao and Mitra (1971, Lemma 5.6.1),

$$\text{rank}(\underset{\sim}{A}) = \text{tr}\underset{\sim}{A}^2, \tag{3.1}$$

since $\underset{\sim}{A}^2$ is now idempotent, and has rank equal to $\text{tr}\underset{\sim}{A}^2 = \text{rank}(\underset{\sim}{A}^2) \geq \text{rank}(\underset{\sim}{A}^3) = \text{rank}(\underset{\sim}{A}) \geq \text{rank}(\underset{\sim}{A}^2)$. Thus

$$\sum \text{rank}(\underset{\sim}{A}_i) = \sum\text{tr}\underset{\sim}{A}_i^2 = \text{tr}\sum\underset{\sim}{A}_i^2 = \text{tr}\underset{\sim}{A}^2 = \text{rank}(\underset{\sim}{A})$$

when (a"), (b"), (c") hold. Notice that we have not supposed that $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_k$ are symmetric; the equality (3.1) holds even when $\underset{\sim}{A}$ is not symmetric.

As Khatri (1977, p. 88) has pointed out, (c") and (d") need not imply (a") and (b") even if the $\underset{\sim}{A}_i$'s are symmetric; e.g., if

$$\underset{\sim}{A}_1 = \frac{1}{3}\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}, \quad \underset{\sim}{A}_2 = -\frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \underset{\sim}{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

then $\text{rank}(\underset{\sim}{A}_1) + \text{rank}(\underset{\sim}{A}_2) = 1 + 1 = 2 = \text{rank}(\underset{\sim}{A})$ and $\underset{\sim}{A}^3 = \underset{\sim}{A}$, but $\underset{\sim}{A}_1^3 \neq \underset{\sim}{A}_1$, $\underset{\sim}{A}_2^3 \neq \underset{\sim}{A}_2$, $\underset{\sim}{A}_1\underset{\sim}{A}_2 \neq \underset{\sim}{0}$. It is, therefore, of interest to see what extra

condition could be added to (c") and (d") so as to imply (a") and (b"). Khatri (1977, Lemma 10) uses

$$\operatorname{rank}(\underset{\sim}{A}) = \sum\{\operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{A}^2 + \underset{\sim}{A})] + \operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{A}^2 - \underset{\sim}{A})]\}, \tag{3.2}$$

which is rather complicated. We may simplify (3.2) in various ways. To do this we first note that

$$\underset{\sim}{A}^3 = \underset{\sim}{A} \leftrightarrow \underset{\sim}{A} = \underset{\sim}{A}^-, \tag{3.3}$$

cf. Graybill (1969, Theorem 12.4.1), Rao and Mitra (1971, Lemma 5.6.2). Thus Theorem 1.2 implies that (c"), (d") are equivalent to

$$\underset{\sim}{A}_i\underset{\sim}{A}\underset{\sim}{A}_i = \underset{\sim}{A}_i, \qquad i=1,\ldots,k, \tag{3.4}$$

and

$$\underset{\sim}{A}_i\underset{\sim}{A}\underset{\sim}{A}_j = \underset{\sim}{0} \qquad \text{for all } i\neq j. \tag{3.5}$$

Summing (3.5) over all $j\neq i$ and adding to (3.4) yields

$$\underset{\sim}{A}_i\underset{\sim}{A}^2 = \underset{\sim}{A}_i, \qquad i=1,\ldots,k. \tag{3.6}$$

Hence under (c") and (d") the condition (3.2) is equivalent to

$$\operatorname{rank}(\underset{\sim}{A}) = \sum\{\operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{I} + \underset{\sim}{A})] + \operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{I} - \underset{\sim}{A})]\}, \tag{3.7}$$

which is a little simpler than (3.2). But (3.7) implies

$$\begin{aligned}\operatorname{rank}(\underset{\sim}{A}) &\geq \sum\operatorname{rank}(2\underset{\sim}{A}_i) = \sum\operatorname{rank}(\underset{\sim}{A}_i) \\ &= \operatorname{rank}(\underset{\sim}{A})\end{aligned} \tag{3.8}$$

when (d") holds. Thus equality holds throughout (3.8) and so (3.7) implies

$$\operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{I} + \underset{\sim}{A})] + \operatorname{rank}[\underset{\sim}{A}_i(\underset{\sim}{I} - \underset{\sim}{A})], \quad i=1,\ldots,k. \tag{3.9}$$

Summing (3.9) and using (d") yields (3.7).

Some motivation for the condition (3.9), and hence also for the equivalent conditions (3.2) and (3.7), may be obtained from the following characterization of a tripotent matrix, extending Lemma 5.6.6 in Rao and Mitra (1971, p. 114):

LEMMA 3.1. *Let* $\underset{\sim}{A}$ *be a square matrix, not necessarily symmetric. Then* $\underset{\sim}{A}^3 = \underset{\sim}{A}$ *if and only if*

$$\operatorname{rank}(\underset{\sim}{A}) = \operatorname{rank}(\underset{\sim}{A}+\underset{\sim}{A}^2) + \operatorname{rank}(\underset{\sim}{A}-\underset{\sim}{A}^2). \tag{3.10}$$

Proof. We use Theorem 1.2 with $\underset{\sim}{A}_1 = \underset{\sim}{A} + \underset{\sim}{A}^2$, $\underset{\sim}{A}_2 = \underset{\sim}{A} - \underset{\sim}{A}^2$, and $\underset{\sim}{A}_1 + \underset{\sim}{A}_2 = 2\underset{\sim}{A}$. If $\underset{\sim}{A}^3 = \underset{\sim}{A}$ then $\frac{1}{2}\underset{\sim}{A} = (2\underset{\sim}{A})^-$ and (3.10) follows from (1.9) since

$$\begin{aligned}(\underset{\sim}{A}+\underset{\sim}{A}^2)\,\tfrac{1}{2}\underset{\sim}{A}(\underset{\sim}{A}+\underset{\sim}{A}^2) &= \tfrac{1}{2}\underset{\sim}{A}^3 + \underset{\sim}{A}^4 + \tfrac{1}{2}\underset{\sim}{A}^5 \\ &= \underset{\sim}{A} + \underset{\sim}{A}^2,\end{aligned}$$

and similarly

$$(\underset{\sim}{A}-\underset{\sim}{A}^2)\,\tfrac{1}{2}\underset{\sim}{A}(\underset{\sim}{A}-\underset{\sim}{A}^2) = \underset{\sim}{A} - \underset{\sim}{A}^2.$$

To go the other way we use (1.11). Then (3.10) implies

$$\underset{\sim}{0} = (\underset{\sim}{A}+\underset{\sim}{A}^2)\,\tfrac{1}{2}\underset{\sim}{A}^-\ (\underset{\sim}{A}-\underset{\sim}{A}^2) = \tfrac{1}{2}\underset{\sim}{A} - \tfrac{1}{2}\underset{\sim}{A}^3$$

and so $\underset{\sim}{A}^3 = \underset{\sim}{A}$ and the proof is complete. Q.E.D.

This suggests using the condition

$$\underset{\sim}{A}_i\underset{\sim}{A} = \underset{\sim}{A}_i^2, \qquad i=1,\ldots,k, \tag{3.11}$$

instead of (3.9), or (3.7) or (3.2). We obtain:

THEOREM 3.1. Let $A_1, \dots, A_k$ be square matrices, not necessarily symmetric, and let $A = \sum A_i$. Then

$$(a'') \qquad A_i^3 = A_i, \qquad i=1,\dots,k,$$

and

$$(b'') \qquad A_i A_j = 0 \qquad \text{for all } i \neq j,$$

hold if and only if

$$(c'') \qquad A^3 = A,$$

$$(d'') \qquad \sum \operatorname{rank}(A_i) = \operatorname{rank}(A),$$

and

$$(e'') \qquad A_i A = A_i^2, \qquad i=1,\dots,k.$$

The condition (e") may be replaced by (3.9), by (3.7), by (3.2), by

$$(e1) \qquad A_i^2 A = A_i \qquad i=1,\dots,k,$$

or by

$$(e2) \qquad A_i A = A A_i, \qquad i=1,\dots,k.$$

Proof. We have already shown that (a"), (b") imply (c"), (d") and hence also (e"), (3.9), (3.7), (3.2), (e1) and (e2). To go the other way let (c"), (d") hold. Then (3.4) and (3.5) are true. Substituting (e") yields (a") and $A_i^2 A_j = 0$ for all $i \neq j$; premultiplying by A_i yields (b"). We have shown that when (c"), (d") hold, then (3.9), (3.7) and (3.2) are equivalent. To see that (a"), (b") are implied we use Theorem 1.2 with the A_i's replaced by the $A_i(I + A)$ and the $A_i(I - A)$ in (3.7), which equation shows them to be rank-additive (the sum is $2A$). Then (1.11) implies that

$$A_i(I + A)(\tfrac{1}{2}A)A_i(I - A) = 0, \tag{3.12}$$

using $\frac{1}{2}A = (2A)^-$. Substituting (3.4) and (3.6) into (3.12) yields

$$(A_i + A_i^2)(I - A) = 0.$$

Postmultiplying by A_j ($j \neq i$) and using (3.5) gives

$$A_i A_j = -A_i^2 A_j. \tag{3.13}$$

However, (1.11) also implies, cf. (3.12),

$$A_i(I - A)(\tfrac{1}{2}A)A_i(I + A) = 0,$$

which leads to

$$A_i A_j = A_i^2 A_j. \tag{3.14}$$

Adding (3.13) and (3.14) yields (b"), and substituting (b") into (3.4) gives (a").

Now let (c"), (d"), (e1) hold. Then (3.4), (3.5) hold and premultiplying (3.5) by A_i yields (b"). Then (3.4) implies (a"). Finally, we let (c") (d") (e2) hold. Then substitution of (e2) into (3.4) yields (e1), and the proof is complete. Q.E.D.

Khatri (1977, Lemma 10) proved the part of Theorem 3.1 with (e") replaced by (3.2). He also claimed that (b"), (c") $\rightarrow$ (a"), (d"), (e"). But this is not so for the same reason that this does not hold in Theorem 1.1; again if we let

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \tag{3.15}$$

then (b"), (c") hold, but (a"), (d") do not. If, however, we add the *condition*

$$\text{(e''')} \quad \operatorname{rank}(A_i^2) = \operatorname{rank}(A_i)$$

to (b"), (c") as was done in Theorem 1.1, then (a"), (d") do follow. From (b"), (c") we have

$$\sum A_i^3 = \sum A_i$$

and so, cf. (2.8)

$$A_i^4 = A_i^2,$$

which implies $A_i^3 = A_i$ using (e''') and Lemma 2.1.

The commutativity condition in Theorem 3.1:

$$\text{(e2)} \qquad A_i A = A A_i, \qquad i=1,\ldots,k,$$

has been used before in a generalization of Cochran's Theorem. Recall the statements.

$$\text{(b'')} \qquad A_i A_j = 0 \qquad \text{for all } i \neq j,$$

$$\text{(d'')} \quad \sum \operatorname{rank}(A_i) = \operatorname{rank}(A),$$

$$\text{(e''')} \quad \operatorname{rank}(A_i^2) = \operatorname{rank}(A_i), \qquad i=1,\ldots,k.$$

Then Marsaglia (1967, Theorem 3) and Marsaglia and Styan (1974, Theorem 15) proved that

$$\text{(b''), (d'')} \leftrightarrow \text{(d''), (e2)}$$

$$\text{(b''), (e''')} \rightarrow \text{(d''),} \tag{3.16}$$

while Luther (1965, Theorem 1, p. 684) and Taussky (1966, Theorem 2), assuming the $\underset{\sim}{A}_i$'s to be symmetric, proved that

$$(b'') \leftrightarrow (d''), (e2).$$

The condition (e''') in (3.16) cannot be dropped in view of the example (3.15); when the $\underset{\sim}{A}_i$'s are symmetric, however, (e''') is automatically satisfied.

Luther (1965, Theorem 3, p. 689) and Tan (1976, Theorem 2.2) have given versions of Theorem 3.1 when the $\underset{\sim}{A}_i$'s are symmetric. We obtain:

THEOREM 3.2. *Let* $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_k$ *be symmetric matrices and let* $\underset{\sim}{A} = \sum \underset{\sim}{A}_i$. *Then*

$$(a'') \qquad \underset{\sim}{A}_i^3 = \underset{\sim}{A}_i, \qquad i=1,\ldots,k,$$

and

$$(b'') \qquad \underset{\sim}{A}_i\underset{\sim}{A}_j = 0 \qquad \textit{for all } i \neq j$$

hold if and only if

$$(c'') \qquad \underset{\sim}{A}^3 = \underset{\sim}{A},$$

$$(d'') \qquad \sum_{i=1}^{k} \operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{rank}(\underset{\sim}{A}),$$

and

$$(es) \qquad \operatorname{tr}\underset{\sim}{A}\underset{\sim}{A}_i \geq \operatorname{tr}\underset{\sim}{A}_i^2, \qquad i=1,\ldots,k-1.$$

The condition (es) *may be replaced by*

$$(es1) \qquad \operatorname{tr}\underset{\sim}{A}^2 \geq \sum_{i=1}^{k} \operatorname{tr}\underset{\sim}{A}_i^2$$

or by

$$(es2) \qquad \operatorname{rank}(\underset{\sim}{A}_i) \geq \operatorname{tr}\underset{\sim}{A}_i^2, \qquad i=1,\ldots,k-1.$$

The condition (es2) was used by Luther (1965, Theorem 3, p. 689), who also showed that the condition (es) may be replaced by

$$\operatorname{tr}\mathbf{A}_i\mathbf{A}_j \geq 0 \qquad \text{for all } i \neq j; \tag{3.17}$$

summing (3.17) over all $i \neq j$ yields (es1). Luther also considered the condition

$$\mathbf{A}_i^3 = \mathbf{A}_i, \qquad i=1,\ldots,k-1, \tag{3.18}$$

and proved that (3.18), (c"), (d") → (a"), (b"), while Khatri (1977, Lemma 10 and Note 9) showed that (a"), (c"), (d") → (b"). But (a") clearly implies (3.18), which implies (es2) with equality in view of (3.1). Tan (1976, Theorem 2.2) gives a condition which seems to be intended to be (es1) with equality (Tan has an extra ‡ (in his notation) inside the trace on both sides of his condition).

Our proof of Theorem 3.2 uses the following result, cf. Graybill (1969, p. 235).

LEMMA 3.2. *Let* $\mathbf{A}$ *be a square matrix. Then*

$$\operatorname{tr}\mathbf{A}'\mathbf{A} \geq \operatorname{tr}\mathbf{A}^2$$

with equality if and only if $\mathbf{A}$ *is symmetric*.

Proof. The result follows at once from

$$\operatorname{tr}(\mathbf{A}-\mathbf{A}')'(\mathbf{A}-\mathbf{A}') = 2(\operatorname{tr}\mathbf{A}'\mathbf{A} - \operatorname{tr}\mathbf{A}^2) \geq \mathbf{0}. \qquad \text{Q.E.D.}$$

Proof of Theorem 3.2. That (a"), (b") imply (c"), (d"), (es), (es1), (es2) follows from Theorem 3.1. To go the other way, let (c"), (d") hold. Then (3.4) and (3.6) hold, and so

$$\mathrm{tr}\mathbf{A}_i\mathbf{A}^2\mathbf{A}_i = \mathrm{tr}(\mathbf{A}\mathbf{A}_i)'\mathbf{A}\mathbf{A}_i \geq \mathrm{tr}(\mathbf{A}\mathbf{A}_i)^2 = \mathrm{tr}\mathbf{A}\mathbf{A}_i\mathbf{A}\mathbf{A}_i \tag{3.19}$$

becomes

$$\mathrm{tr}\mathbf{A}_i^2 \geq \mathrm{tr}\mathbf{A}\mathbf{A}_i. \tag{3.20}$$

The condition (es) implies equality in (3.20), and hence in (3.19) and so by Lemma 3.2 $\mathbf{A}\mathbf{A}_i = (\mathbf{A}\mathbf{A}_i)' = \mathbf{A}_i\mathbf{A}$, which is condition (e2) of Theorem 3.1, but only for i=1,...,k-1. Substitution in (3.6) yields

$$\mathbf{A}\mathbf{A}_i\mathbf{A}_j = \mathbf{0}, \qquad i=1,\dots,k-1 \text{ and } j\neq i. \tag{3.21}$$

But (3.4) implies that $\mathrm{rank}(\mathbf{A}\mathbf{A}_i) = \mathrm{rank}(\mathbf{A}_i)$ and so by Lemma 2.2 we may cancel the $\mathbf{A}$ in (3.21) to get

$$\mathbf{A}_i\mathbf{A}_j = \mathbf{0}, \qquad i=1,\dots,k-1 \text{ and } j\neq i.$$

Thus

$$\mathbf{A}_i\mathbf{A}_k = \mathbf{0} = \mathbf{A}_k\mathbf{A}_i, \qquad i=1,\dots,k-1,$$

upon transposition and so (b") holds. Substitution in (3.4) yields (a").

Now suppose (c"), (d"), (es1) hold. Then so does (3.20) which we may sum to yield

$$\sum_{i=1}^{k} \mathrm{tr}\mathbf{A}_i^2 \geq \mathrm{tr}\mathbf{A}^2.$$

But (es1) indicates that this inequality goes the other way and so we must have equality, which in turn implies equality in (3.19) and that $\mathbf{A}\mathbf{A}_i$ is symmetric for all i=1,...,k. Thus (a"), (b") are implied as before.

Finally let (c"), (d"), (es2) hold. Then from (3.4) $\mathbf{A}\mathbf{A}_i$ is idempotent and so

$$\mathrm{tr}\mathbf{A}\mathbf{A}_i = \mathrm{rank}(\mathbf{A}\mathbf{A}_i) = \mathrm{rank}(\mathbf{A}_i) \geq \mathrm{tr}\mathbf{A}_i^2, \qquad i=1,\dots,k-1,$$

is condition (es) and our proof is complete. Q.E.D.

We conclude this section with an extension of Theorem 3.1 to r-potent matrices. We define a square matrix $\mathbf{A}$ to be r-potent whenever $\mathbf{A}^r = \mathbf{A}$ for some positive integer $r \geq 2$. As Tan (1975, Lemma 1) has pointed out, the nonzero eigenvalues of an r-potent matrix are the (r-1)th roots of unity. Since a symmetric matrix has only real eigenvalues, a symmetric r-potent matrix must be tripotent even though Tan (1976, p. 608) suggests otherwise.

We obtain:

THEOREM 3.3. Let $\mathbf{A}_1,\ldots,\mathbf{A}_k$ be square matrices, not necessarily symmetric, and let $\mathbf{A} = \sum \mathbf{A}_i$. Let r be a fixed positive integer ≥ 2. Then

(a) $\quad \mathbf{A}_i^r = \mathbf{A}_i, \qquad i=1,\ldots,k$

and

(b) $\quad \mathbf{A}_i\mathbf{A}_j = \mathbf{0} \qquad$ for all $i \neq j$

hold if and only if

(c) $\quad \mathbf{A}^r = \mathbf{A},$

(d) $\quad \operatorname{rank}(\mathbf{A}_i) = \operatorname{rank}(\mathbf{A}),$

and

$(e)_r \quad \mathbf{A}_i\mathbf{A}^{r-2} = \mathbf{A}_i^{r-1}, \qquad i=1,\ldots,k.$

The condition $(e)_r$ may be replaced by

$(e1)_r \quad \mathbf{A}_i^2\mathbf{A}^{r-2} = \mathbf{A}_i, \qquad i=1,\ldots,k,$

or by

$(e2)_r \quad \mathbf{A}_i\mathbf{A}^{r-2} = \mathbf{A}^{r-2}\mathbf{A}_i, \qquad i=1,\ldots,k.$

Tan (1975, Theorem 2.1) suggested that (c), (d) $\rightarrow$ (a), (b) but, cf. Khatri (1976), seems to have realized that this is not true (Tan, 1976). When $r = 3$ the conditions $(e)_r$, $(e1)_r$, $(e2)_r$ become the conditions (e), (e1), (e2), respectively, of Theorem 3.1. When $r = 2$ the conditions $(e)_r$, $(e2)_r$ are automatically satisfied and Theorem 3.3 becomes part of Theorem 1.1. The condition $(e1)_r$, however, when $r = 2$ becomes $\underset{\sim}{A}_i^2 = \underset{\sim}{A}_i$, or (a), and so $(e1)_r$ may be too strong an extra condition to require that (c), (d) $\rightarrow$ (a), (b) in Theorem 3.3. Under (c), (d), however, the condition $(e1)_r$ is equivalent to the commutativity condition

$$\underset{\sim}{A}_i(\underset{\sim}{A}_i\underset{\sim}{A}^{r-2}) = (\underset{\sim}{A}_i\underset{\sim}{A}^{r-2})\underset{\sim}{A}_i, \qquad i=1,\ldots,k, \tag{3.22}$$

which is in the same spirit as the condition $(e2)_r$. When $r = 2$ the condition (3.22) is automatically satisfied.

To prove that the conditions (3.22) and $(e1)_r$ are equivalent when (c), (d) hold we note first that $\underset{\sim}{A}^{r-2} = \underset{\sim}{A}^-$ when $\underset{\sim}{A}$ is r-potent. Then, cf. (3.4)-(3.6), we see that (c), (d) are equivalent to

$$\underset{\sim}{A}_i\underset{\sim}{A}^{r-2}\underset{\sim}{A}_i = \underset{\sim}{A}_i, \qquad i=1,\ldots,k, \tag{3.23}$$

$$\underset{\sim}{A}_i\underset{\sim}{A}^{r-2}\underset{\sim}{A}_j = \underset{\sim}{0} \qquad \text{for all } i\neq j, \tag{3.24}$$

$$\underset{\sim}{A}_i\underset{\sim}{A}^{r-1} = \underset{\sim}{A}_i, \qquad i=1,\ldots,k, \tag{3.25}$$

and (3.23) shows that $(3.22) \leftrightarrow (e1)_r$.

<u>Proof of Theorem 3.3</u>. Let (a), (b) hold. Then so does (c), and

$$(\underset{\sim}{A}_i^{r-1})^2 = \underset{\sim}{A}_i^{r+r-2} = \underset{\sim}{A}_i^{r-1}$$

is idempotent and so

$$\operatorname{rank}(\underset{\sim}{A}_i) = \operatorname{tr}\underset{\sim}{A}_i^{r-1},$$

cf. (3.1). Hence (a), (b), (c) imply

$$\sum \operatorname{rank}(\underset{\sim}{A}_i) = \sum \operatorname{tr}\underset{\sim}{A}_i^{r-1} = \operatorname{tr} \sum \underset{\sim}{A}_i^{r-1} = \operatorname{tr}\underset{\sim}{A}^{r-1} = \operatorname{rank}(\underset{\sim}{A}),$$

which is (d). Then (b) implies $(e)_r$ and $(e2)_r$ and turns $(e1)_r$ into (a). To go the other way, let (c), (d), $(e)_r$ hold. Then so do (3.23), (3.24). Substitution of $(e)_r$ into (3.23) yields (a), while substitution of $(e)_r$ into (3.24) yields $\underset{\sim}{A}_i^{r-1}\underset{\sim}{A}_j = \underset{\sim}{0} = \underset{\sim}{A}_i\underset{\sim}{A}_j$ upon premultiplication by $\underset{\sim}{A}_i$ and substituting (a).

Now let (c), (d), $(e1)_r$ hold. Then (3.23), (3.24) hold and postmultiplying $(e1)_r$ by $\underset{\sim}{A}_j$ $(j \neq i)$ yields (b) by substituting (3.24). Then (a) follows from (3.23) by use of (b). Finally, suppose that (c), (d), $(e2)_r$ hold. Premultiplying $(e2)_r$ by $\underset{\sim}{A}_i$ and substituting (3.23) yields $(e1)_r$ and so our proof is complete. Q.E.D.

Tan (1975, Theorem 2.1) also suggested that (b), (c) and

$$\operatorname{rank}(\underset{\sim}{A}_i^{r-1}) = \operatorname{rank}(\underset{\sim}{A}_i^{2(r-1)}), \qquad i=1,\ldots,k, \tag{3.26}$$

imply (a), (d), but withdrew this, cf. Tan (1976, p. 608). It is straightforward, however, to see that (b), (c) imply

$$\sum \underset{\sim}{A}_i^r = \underset{\sim}{A}_i$$

and hence

$$\underset{\sim}{A}_i^{r+1} = \underset{\sim}{A}_i^2. \tag{3.27}$$

The extra condition

$$\operatorname{rank}(\underset{\sim}{A}_i^2) = \operatorname{rank}(\underset{\sim}{A}_i),$$

cf. (e''') , applied to (3.27) then yields (a) in view of the rank cancellation rule Lemma 2.1. The extra condition (3.26) is, however, not sufficient (unless $r=2$), as is seen from the counter-example provided by (3.15).

4. Statistical Applications

The analysis of variance involves the decomposition of a sum of squares of observations into quadratic forms. In classical cases these quadratic forms are independently distributed according to χ^2-distributions. Then ratios of them are proportional to F-statistics. Cochran's Theorem provides an algebraic method of verifying the necessary properties of the quadratic forms to justify the F-tests.

As indicated in Lemma 1.1, when x has the distribution $\underline{\underline{N}}(0, I)$ then $A^2 = A$ implies $x'Ax$ has the χ^2-distribution with degrees of freedom equal to the number of unit eigenvalues of A, the other eigenvalues being 0. Lemma 1.2 states that $AB = 0$ implies independence of $x'Ax$ and $x'Bx$ because the joint characteristic function when $x \sim \underline{\underline{N}}(0, I)$ is, cf. (1.3),

$$\mathcal{E}\, e^{\frac{1}{2}isx'Ax + \frac{1}{2}itx'Bx} = |I - isA - itB|^{-\frac{1}{2}p} = |I - isA|^{-\frac{1}{2}p}|I - itB|^{-\frac{1}{2}p}.$$

As an example, consider the one-way analysis of variance. Let $y_{i\alpha}$ be normally distributed according to $\underline{\underline{N}}(\mu_i, \sigma^2)$, $i=1,\dots,m$, $\alpha = 1,\dots,n$, and suppose the mn variables are independent. Under the typical null hypothesis H: $\mu_1 = \dots = \mu_m = \mu$, say, the exponent of the normal

distribution is $\sum_{i=1}^{m} \sum_{j=1}^{n} (y_{i\alpha} - \mu).^2$ Let

$$q_1 = n \sum_{i=1}^{m} (\bar{y}_i - \bar{y})^2 = n \sum_{i=1}^{m} \bar{y}_i^2 - mn\bar{y}^2,$$

$$q_2 = \sum_{i=1}^{m} \sum_{\alpha=1}^{n} (y_{i\alpha} - \bar{y}_i)^2 = \sum_{i=1}^{m} \sum_{\alpha=1}^{n} y_{i\alpha}^2 - n \sum_{i=1}^{m} \bar{y}_i^2 ,$$

$$q_3 = mn\,\bar{y}^2,$$

where $\bar{y}_i = \sum_{\alpha=1}^{n} y_{i\alpha}/N$ and $\bar{y} = \sum_{i=1}^{m} \bar{y}_i/m$. Let $\underset{\sim}{y}^{(i)} = (y_{i1},\ldots,y_{in})'$, $\underset{\sim}{y} = (\underset{\sim}{y}^{(1)'},\ldots,\underset{\sim}{y}^{(m)'})'$,

$$\underset{\sim}{A}_1 = \frac{1}{n}(\underset{\sim}{I}_m - \frac{1}{m}\underset{\sim}{\varepsilon}_m\underset{\sim}{\varepsilon}_m') \otimes \underset{\sim}{\varepsilon}_n\underset{\sim}{\varepsilon}_n',$$

$$\underset{\sim}{A}_2 = \underset{\sim}{I}_m \otimes (\underset{\sim}{I}_n - \frac{1}{n}\underset{\sim}{\varepsilon}_n\underset{\sim}{\varepsilon}_n'),$$

$$\underset{\sim}{A}_3 = \frac{1}{mn}\,\underset{\sim}{\varepsilon}_m\underset{\sim}{\varepsilon}_m' \otimes \underset{\sim}{\varepsilon}_n\underset{\sim}{\varepsilon}_n',$$

where $\underset{\sim}{\varepsilon}_n = (1,\ldots,1)'$ of n components and $\underset{\sim}{\varepsilon}_m = (1,\ldots,1)'$ of m components. Then $q_i = \underset{\sim}{y}'\underset{\sim}{A}_i\underset{\sim}{y}$. We easily verify that $\Sigma \underset{\sim}{A}_i = \underset{\sim}{I}_{mn}$, $\text{rank}(\underset{\sim}{A}_1) = m-1$, $\text{rank}(\underset{\sim}{A}_2) = m(n-1)$, and $\text{rank}(\underset{\sim}{A}_3) = 1$. Then (a) and (b) hold. (Of course, in the simple example above the conditions could be verified directly.) By Lemmas 1.1 and 1.2 the quadratic forms are independently distributed as χ^2's, the last being noncentral.

The multivariate analogue of the χ^2-distribution is the Wishart distribution. If $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_N$ are independently distributed, each according to $\underline{\underline{N}}(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, then the distribution of $\underset{\sim}{S} = \sum_{\alpha=1}^{N} \underset{\sim}{Y}_\alpha\underset{\sim}{Y}_\alpha'$ is known as the Wishart distribution. (Cf. e.g., Chapter 7 of Anderson (1958).) If $q_1,\ldots,q_k$ have independent χ^2-distributions when the dimensionality of $\underset{\sim}{Y}_\alpha$ is 1,

then $\underset{\sim}{Q}_1 = \sum_{\alpha,\beta=1}^{N} a_{\alpha\beta}^{(1)} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\beta', \ldots, \underset{\sim}{Q}_k = \sum_{\alpha,\beta=1}^{N} a_{\alpha\beta}^{(k)} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\beta'$ have independent Wishart distributions; here $\underset{\sim}{A}_i = (a_{\alpha\beta}^{(i)})$, $i=1,\ldots,k$. Cochran's Theorem is correspondingly useful in multivariate analysis of variance.

It should be noted that when $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_k$ are symmetric several proofs show that there exists an orthogonal matrix that simultaneously diagonalizes $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_k$, the resulting diagonal matrices have 0's and 1's as diagonal elements, and the 1's in the transformed $\underset{\sim}{A}_i$ correspond to 0's in the transformed $\underset{\sim}{A}_j$, $j\neq i$. Cf. (2.3)-(2.4).

If $\underset{\sim}{A}^3 = \underset{\sim}{A}$, the eigenvalues of $\underset{\sim}{A}$ are 1, -1, and 0. Hence $\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x}$ for $\underset{\sim}{x} \sim \underline{\underline{N}}(\underset{\sim}{0}, \underset{\sim}{I})$ has the distribution of $\chi_1^2 - \chi_2^2$, where χ_1^2 and χ_2^2 are independent, the number of degrees of freedom of χ_1^2 is the number of eigenvalues equal to 1 and the number of degrees of freedom of χ_2^2 is the number of eigenvalues equal to -1.

Components of variance are often estimated as differences of quadratic forms. Let $y_{i\alpha} = \mu + u_i + v_{i\alpha}$, $\alpha=1,\ldots,n$, $i=1,\ldots,m$, where μ is an unobservable constant and the unobservable u_i's and $v_{i\alpha}$'s are independently normally distributed with means 0 and variances $\mathcal{E}u_i^2 = \sigma_u^2$ and $\mathcal{E}v_{i\alpha}^2 = \sigma_v^2$. Then for q_1 and q_2 as defined above

$$\mathcal{E}q_1 = (m-1)(n\sigma_u^2 + \sigma_v^2),$$

$$\mathcal{E}q_2 = (mn-m)\sigma_v^2.$$

Thus $q_1\ (m-1) - q_2/(mn-m)$ is an unbiased estimator of $n\sigma_u^2$. Other differences of quadratic forms arise in other designs.

Press (1966) has given the distribution of an arbitrary quadratic form, which is a linear combination of χ^2's with possibly negative coefficients. Let $Z = \alpha\chi_1^2 - \beta\chi_2^2$, where χ_1^2 and χ_2^2 are independently distributed as χ^2-variables with m and n degrees of freedom, respectively,

and $\alpha > 0$, $\beta > 0$. The density of Z is

$$[K/\Gamma(\tfrac{1}{2}m)]t^{\frac{1}{2}(m+n)-1} e^{-z/2\alpha}\psi[\tfrac{1}{2}n, \tfrac{1}{2}(m+n); t(\alpha+\beta)/2\alpha\beta], \quad t \geq 0,$$

$$[K/\Gamma(\tfrac{1}{2}n)](-t)^{\frac{1}{2}(m+n)-1} e^{t/2\beta}\psi[\tfrac{1}{2}m, \tfrac{1}{2}(m+n); -t(\alpha+\beta)/2\alpha+\beta], \quad t \leq 0,$$

where $K^{-1} = 2^{\frac{1}{2}(m+n)} \alpha^{\frac{1}{2}m}\beta^{\frac{1}{2}n}$ and

$$\psi(a,b,x) = \frac{\Gamma(1-b)}{\Gamma(1+a-b)}\ {}_1F_1(a,b;x) + \frac{\Gamma(b-1)}{\Gamma(a)}\ x^{1-b}\ {}_1F_1(1+a-b,2-b;x),$$

and ${}_1F_1(a,b;x)$ is the confluent hypergeometric function. Robinson (1965) gave a similar result for $\alpha=\beta=1$. In the special case of equal degrees of freedom $(n=m)$ Pearson, Stouffer and David (1932) gave the density of $Z = \chi_1^2 - \chi_2^2$ as

$$\frac{z^{p-\frac{1}{2}}\ K_{n-\frac{1}{2}}(2z)}{2\sqrt{\pi}\ \Gamma(n)},$$

where $K_r(x)$ is the Bessel function of second order and imaginary argument.

In Theorem 3.2 (a")indicates that $q_i = x'A_ix$ is distributed as the difference of two χ^2-variables if $x \sim N(0, I)$ and (b")states that q_i and q_j are independent. Then (c")and (d")and either (es), (es1), or (es2) are conditions implying (a")and (b"). In most cases (c") is easily verified and (d")is as in Section 1. Each of (es), (es1), and (es2) require computation of trA_i^2, $i=1,\ldots,k-1$, and (es1) needs also trA_k^2. Of the left-hand sides, trA^2 may be easiest to compute.

REFERENCES

AITKEN, A. C. (1950). On the statistical independence of quadratic forms in normal variates. Biometrika 37 93-96.

ANDERSON, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

ANDERSON, T. W., DAS GUPTA, S. AND STYAN, G. P. H. (1972). A Bibliography of Multivariate Statistical Analysis. Oliver & Boyd, Edinburgh, Scotland. (Reprinted 1977, Krieger, Huntington, N.Y.)

BANERJEE, K. S. AND NAGASE, G. (1976). A note on the generalization of Cochran's theorem. Comm. Statist. A--Theory Methods 5 837-842.

CARPENTER, O. (1950). Note on the extension of Craig's Theorem to non-central variates. Ann. Math. Statist. 21 455-457.

CHIPMAN, J. S., AND RAO, M. M. (1964). Projections, generalized inverses, and quadratic forms. J. Math. Anal. Appl. 9 1-11.

COCHRAN, W. G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Proc. Cambridge Philos. Soc. 30 178-191.

CRAIG, A. T. (1938). On the independence of certain estimates of variance. Ann. Math. Statist. 9 48-55.

CRAIG, A. T. (1943). Note on the independence of certain quadratic forms. Ann. Math. Statist. 14 195-197.

FISHER, R. A. (1925). Applications of "Student's" distribution. Metron 5 90-104.

GRAYBILL, F. A. (1969). Introduction to Matrices with Applications in Statistics. Wadsworth, Belmont, Calif.

GRAYBILL, F. A. AND MARSAGLIA, G. (1957). Idempotent matrices and quadratic forms in the general linear hypothesis. Ann. Math. Statist. 28 678-686.

HARTWIG, R. E. (1980). A note on rank-additivity. Linear and Multilinear Algebra, in press.

JAMES, G. S. (1952). Notes on a theorem of Cochran. Proc. Cambridge Philos. Soc. 48 443-446.

KHATRI, C. G. (1968). Some results for the singular normal multivariate regression models. Sankhyā Ser. A 30 267-280.

KHATRI, C. G. (1976). Review of Tan (1975). Math. Reviews 51 #2159.

KHATRI, C. G. (1977). Quadratic forms and extension of Cochran's theorem to normal vector variables. Multivariate Analysis--IV (P. R. Krishnaiah, ed.). North-Holland, Amsterdam, 79-94.

KRAFFT, O. (1978). Lineare statistische Modelle und optimale Versuchspläne. Vandenhoeck & Ruprecht, Göttingen.

LOYNES, R. M. (1966). On idempotent matrices. Ann. Math. Statist. 37 295-296.

LUTHER, N. Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms. Ann. Math. Statist. 36 683-690.

MADOW, W. G. (1940). The distribution of quadratic forms in non-central normal random variables. Ann. Math. Statist. 11 100-103.

MARSAGLIA, G. (1967). Bounds on the rank of the sum of matrices. Trans. Fourth Prague Conf. Information Theory, Statist. Decision Functions, Random Processes (Prague, Aug. 31-Sept. 11, 1965), Czech. Acad. Sci. 455-462.

MARSAGLIA, G. AND STYAN, G. P. H. (1974). Equalities and inequalities for ranks of matrices. Linear and Multilinear Algebra 2 269-292.

OGASAWARA, T. AND TAKAHASHI, M. (1951). Independence of quadratic quantities in a normal system. J. Sci. Hiroshima Univ. Ser. A 15 1-9.

OGAWA, J. (1946). On the independence of statistics of quadratic forms. (In Japanese.) Res. Mem. Inst. Statist. Math. Tokyo 2 98-111.

OGAWA, J. (1947). On the independence of statistics of quadratic forms. (Translation of Ogawa, 1946.) Res. Mem. Inst. Statist. Math. Tokyo 3 137-151.

OGAWA, J. (1948). On the independence of statistics between linear forms, quadratic forms and bilinear forms from normal distributions. (In Japanese.) Res. Mem. Inst. Statist. Math. Tokyo 4 1-40.

OGAWA, J. (1949). On the independence of bilinear and quadratic forms of a random sample from a normal population. (Translation of Ogawa, 1948.) Ann. Inst. Statist. Math. Tokyo 1 83-108.

PEARSON, K., STOUFFER, S. A. AND DAVID, F. N. (1932). Further applications in statistics of the $T_m(x)$ Bessel function. Biometrika 24 293-350.

PRESS, S. J. (1966). Linear combinations of non-central chi-square variates. Ann. Math. Statist. 37 480-487.

RAO, C. R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. J. Roy. Statist. Soc. Ser. B 24 152-158.

RAO, C. R. (1973). Linear Statistical Inference and Its Applications, 2nd ed. Wiley, New York.

RAO, C. R. AND MITRA, S. K. (1971). Generalized Inverse of Matrices and its Applications.

ROBINSON, J. (1965). The distribution of a general quadratic form in normal variates. Austral. J. Statist. 7 110-114.

SAKAMOTO, H. (1944). On the independence of (two) statistics. [Lecture at the annual Mathematics-Physics Meeting, July 19, 1943.] (In Japanese.) Res. Mem. Inst. Statist. Math. Tokyo 1 (9) 1-25.

SCAROWSKY, I. (1973). Quadratic Forms in Normal Variables. M.Sc. thesis, Dept. Math., McGill Univ.

SEARLE, S. R. (1971). Linear Models. Wiley, New York.

STYAN, G. P. H. (1970). Notes on the distribution of quadratic forms in singular normal variables. Biometrika 57 567-572.

TAN, W. Y. (1975). Some matrix results and extensions of Cochran's theorem. SIAM J. Appl. Math. 28 547-554.

TAN, W. Y. (1976). Errata: Some matrix results and extensions of Cochran's theory. SIAM J. Appl. Math. 30 608-610.

TAN, W. Y. (1977). On the distribution of quadratic forms in normal random variables. Canad. J. Statist. 5 241-250.

TAUSSKY, O. (1966). Remarks on a matrix theorem arising in statistics. Monatsh. Math. 70 461-464.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

26. "The Recurrence Classification of Risk and Storage Processes," J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process," T. W. Anderson and Raul P. Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures," Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity," Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.

40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike, August, 1979.

41. "Approximating Conditional Moments of the Multivariate Normal Distribution," Joseph G. Deken, December 1979.

42. "Methods and Applications of Time Series Analysis - Part I: Regression, Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla, July 1980.

43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson and George P.H. Styan, August, 1980.

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.

40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike, August, 1979.

41. "Approximating Conditional Moments of the Multivariate Normal Distribution," Joseph G. Deken, December 1979.

42. "Methods and Applications of Time Series Analysis - Part I: Regression, Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla, July 1980.

43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson and George P.H. Styan, August, 1980.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 43	2. GOVT ACCESSION NO. AD-A092	3. RECIPIENT'S CATALOG NUMBER 636
4. TITLE (and Subtitle) COCHRAN'S THEOREM, RANK ADDITIVITY, AND TRIPOTENT MATRICES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 T. W. Anderson ~~and~~ George P. H. Styan 15		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217	11	12. REPORT DATE AUGUST 1980
		13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) TR-43		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cochran's theorem, rank additivity, tripotent matrices, chi-square distributions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

Let $A_1, \ldots, A_k$ be symmetric matrices and $A = \sum A_i$. A matrix version of Cochran's Theorem is that (a) $A_i^2 = A_i$, $i=1,\ldots,k$, and (b) $A_iA_j = 0 \; \forall \, i \neq j$, are necessary and sufficient conditions for (d) $\sum \text{rank}(A_i) = \text{rank}(A)$ whenever (c) $A = I$. This paper reviews extensions of the theorem and its statistical interpretations in the literature, presents various proofs of the above theorem, and obtains some generalizations. In particular, (c) above is replaced by $A^2 = A$ and the condition of symmetry is deleted. The relations with (e) $\text{rank}(A_i^2) = \text{rank}(A_i)$, $i=1,\ldots,k$, are explored. Another theorem covers the case of matrices not necessarily square. A is "tripotent" if $A^3 = A$. Then (a') $A_i^3 = A_i$, $i=1,\ldots,k$, and (b') are necessary and sufficient conditions for (c') $A^3 = A$, (d), and one further condition such as (e') $A_iA = A_i^2$, $i=1,\ldots,k$. Variations and statistical applications are treated. Tripotent is replaced by r-potent ($A^r = A$) for $r > 3$.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)